DOCUMENT RESUME

ED 079 372 TM 002 971

AUTHOR	Doppelt, Jerome E.; Bennett, George K.
TITLE	Testing Job Applicants from Disadvantaged Groups.
INSTITUTION	Psychological Corp., New York, N.Y.
REPORT NO	Bull-57
PUB DATE	May 67
NOTE	5p.
AVAILABLE FROM	Test Service Bulletin, The Psychological Corporation, 304 East 45th Street, New York, N.Y. 10017 (free)
JOURNAL CIT	Test Service Bulletin; n57 p1-5 May 1967
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Bulletins; *Disadvantaged Groups; *Equal Opportunities (Jobs); Job Applicants; *Personnel Selection; Predictive Ability (Testing); Racial Discrimination; *Test Bias; *Testing; Testing Problems; Test Reliability

ABSTRACT

Companies have been accused of using tests as a means of maintaining unfair discrimination against groups which have already suffered from many forms of discrimination. It does not follow, however, that tests themselves merit condemnation. Employers who use tests do so because the information furnished by tests is valuable to them in making hiring decisions. Although testing is not free from defect or beyond criticism, appropriately chosen and properly administered ability tests are superior to most available alternatives. Some of the problems stemming from the testing of disadvantaged groups, such as reduced reliability due to the compression of score distributions, are of a technical nature and are unrelated to discrimination. Many of the issues have both psychometric and social aspects. Some current testing procedures should be changed to reduce the fear and hostility that may be engendered by materials felt to be biased or unfair. More attention must be devoted to research with actual score and performance data in order to improve the predictive efficiency of tests. But discrimination in the world of work will not be resolved by improved psychometrics alone. The use of tests in employee selection is inherently a friendly rather than a hostile act to those who come to the job market from backgrounds of limited opportunity. The rejection of measuring instruments which register the consequence of cultural deprivation is merely a modern version of killing the messenger who brings bad news. (Author/KM)

ED 079372

THE PSYCHOLOGICAL CORPORATION

Test Service Bulletin

THE PSYCHOLOGICAL CORPORATION
GEORGE K. BENNETT, *President*

May, 1967
No. 57

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TESTING JOB APPLICANTS FROM DISADVANTAGED GROUPS

JEROME E. DOPPELT AND GEORGE K. BENNETT

TM 002 971

EVERY day people make decisions which affect the lives of other people. When tests play a role in the decision process, considerable feeling may be engendered, not only among those directly concerned but among the "spectators" as well. Those of us who have been seriously concerned with improving the effectiveness of tests have been aware of their limitations and of circumstances which may diminish their usefulness. We have long recognized, for example, that a test of mental ability does not measure native ability but rather it measures the individual's present capability of demonstrating his skills or knowledge. From this demonstration one then makes a decision about the individual's likelihood of success in certain endeavors.

It would be a formidable task to discuss the pros and cons of testing in all the fields in which decisions are, or might be, influenced by tests. We will limit this discussion, therefore, to an area of importance to all of us—the use of tests in the employment of people who are members of "disadvantaged groups."

The term "disadvantaged groups," as used in this discussion, refers to the subgroups in our population who have been victims of educational, cultural, or economic disadvantages. The deprivations suffered by people in such groups are assumed to stem primarily from their membership in the groups. Several observations may be made about the term "disadvantaged groups." First, it seems to be an ethnic or national-origin classification. Negroes and Puerto Ricans are usually included among the disadvantaged; so are Mexican-Americans and Indians. Second, there are subgroups of the white majority who are economically and educationally disadvantaged, but the term is seldom applied to them. Third, the term conveys the impression of a group that is homogeneous with respect to ability and deprivation, in spite of the fact that its members usually show a range in abilities and interests. Of course, this range may be restricted when compared with that of the general population.

Employment decisions require the making of predictions, whether the latter are made from mathematical equations or by intuition. We are concerned, therefore, with problems in predicting job success from the test scores of disadvantaged persons. It has been proposed that procedures which are suitable for a majority of applicants are not appropriate for those who may be described as educationally or culturally deprived. This may possibly be so, but one must guard against oversimplified solutions to a complex problem. Naive solutions which are applied equally to all members of a disadvantaged group (or to different groups of disadvantaged) are likely to be ineffective, and may even create new problems of undesirable discrimination. Let us consider some of the charges leveled against tests in this context, and evaluate some suggested corrective measures.

The contents of this Bulletin are not copyrighted; the articles may be quoted or reprinted without formality other than the customary acknowledgement of the Test Service Bulletin of THE PSYCHOLOGICAL CORPORATION *as the source. There is no charge for the Test Service Bulletin. Copies of any or all will be sent on request.*

Published from time to time by the Test Division of THE PSYCHOLOGICAL CORPORATION *in the interest of promoting greater understanding of the principles and techniques of mental measurement and its applications in guidance, personnel work, and clinical psychology, and for announcing new publications of interest. Address communications to*

304 East 45th Street
New York, N. Y. 10017

Staff of the Test Division

ALEXANDER G. WESMAN, *Director*

JEROME E. DOPPELT, *Associate Director*

JAMES H. RICKS, JR., *Associate Director*

DOROTHY M. CLENDENEN, *Assistant Director*

DAVID O. HERMAN, *Assistant Director*

DAVID H. HALL, *Coordinator of Sales Service and Production*

ESTHER R. HOLLIS, *Manager, Advisory Service*

MARILYN C. LEE, *Senior Research Associate*

DOROTHY M. CLENDENEN, *Editor*

We are grateful to one of our readers, Dr. C. B. Conway, Director, Division of Tests and Standards, Dept. of Education, Victoria, B.C., for writing that the author of the limerick quoted in TSB No. 56 is the late Arthur H. Reginald Buller, a botanist who was once president of the Royal Society of Canada.

Other Test Service Bulletins that may be of interest to you:

No. 37 – How Effective Are Your Tests?
No. 45 – Better Than Chance

Some of the Charges

ANXIETY IN THE TESTING SITUATION. A frequent criticism is that many existing tests do not adequately evaluate the capabilities of members of disadvantaged groups. It is pointed out that the disadvantaged may score poorly because of anxiety about the testing situation and because of low motivation. It is also felt that examiners from the "advantaged" population tend to inhibit the performance of the disadvantaged. This is more likely to be a problem when individual tests rather than group tests are administered, and when the examinees are children rather than job applicants, but some feel it is still a problem when testing in employment offices.

UNFAIRNESS OF CONTENT. It is further maintained that most existing tests, especially verbal measures, emphasize middle-class concepts and information and are therefore unfair to those who have not been exposed to middle-class cultural and educational influences. Consequently, the low test scores which are earned are not indicative of the "true" abilities of the disadvantaged. Predictions of job success made from such scores are therefore held to be inaccurate.

IMPROPER INTERPRETATION OF SCORES. There is the contention that scores do not have the same meaning for the disadvantaged that they do for the advantaged. Both the pattern of scores (when a battery is used) and the level of performance (on even a single test) require interpretation in the light of the background of the examinee. It is argued, however, that for interpreting the scores of disadvantaged persons the usual kinds of background evaluation are not enough. Test scores of the disadvantaged, it is held, should be compared only with test scores of others similarly disadvantaged. In this way, a score which would be considered mediocre, according to norms based on a national sample, might be considered superior when compared with the performance of a particular deprived group.

LACK OF RELEVANCE. It is charged that test items are often not related to the work required on the job for which the applicant is being considered, and that even where relationships can be shown between test scores and job success, there is no need to eliminate low-scoring disadvantaged people since they can be taught the necessary skills and knowledge in a training period after hiring. In addition, some critics feel that bias against disadvantaged groups frequently enters into the performance ratings which serve as the criteria for validating the tests, and this artificially confirms the gloomy predictions made from the scores. Thus, the picture is one of a self-fulfilling prophecy of doom for the disadvantaged.

Meeting the Criticisms

In response to these charges, reasonable people generally offer suggestions which they believe will make the tests "fairer" rather than demand that all testing cease. Eliminating tests entirely would remove from the hiring process an approach which is potentially more objective, more color-blind, and more susceptible to verification as to effectiveness than most of the other techniques used in the employment procedure. As John W. Gardner, the present Secretary of Health, Education, and Welfare, wrote in the book *Excellence*[1] in reference to testing in schools: "The tests couldn't see whether the youngster was in rags or in tweeds, and they couldn't hear the accents of the slum. The tests revealed intellectual gifts at every level of the population." More is to be gained from the proper use of tests than from their exclusion; accordingly, some of the corrective measures that have been offered should be carefully scrutinized.

Methods of Alleviating Anxiety

To allay test anxiety it has been suggested that if an examinee has failed, he be permitted to come back for one or more retests. It is alleged that this procedure would provide the practice in taking tests which is so lacking among the disadvantaged.

Where the test requires demonstration of proficiency, as in typewriting, the test is a sample of the actual skill required for the job, and practice by the applicant between testings would only tend to improve his real skill. The administration of a form that may have been given previously would not result in distorted interpretations of his current typing skill, although it is preferable to have alternate forms available for retesting.

With a general ability test, however, permitting applicants to take it again and again, as a means of reducing anxiety, has serious consequences. Retests are sometimes sought by individuals who hope to become familiar with the nature of the specific test and, consequently, to do better on the retest. If the same form of a test is administered more than once within a short time period, practice between testings is likely to result not in improved general ability that would be pertinent to success on the job, but rather in improved scores which may simply be due to practice in remembering answers to specific questions. Furthermore, the examinee may have searched for outside help to provide answers to some items. The validity of scores obtained under such conditions is suspect.

Nevertheless, people who wish to be retested should probably be given the opportunity to take the test or tests again, but any test which is to be given to the same people more than once should have alternate forms. Although the availability of two or three forms of a test is not uncommon, the preparation of a large number of forms is a time-consuming and expensive process. In addition to alternate forms of the *same* test, it would be desirable to have *different* tests, each with several forms, available for testing and retesting. Of course, studies which establish the comparability of scores on the different tests used would be essential.

It is apparent from the discussion above that the successive administration of "real tests" over a short time interval is not always a practical procedure for allaying test anxiety. Rather, it is preferable to provide a practice or demonstration period in which different kinds of tests and answer media are discussed. This kind of activity can alleviate the unfamiliarity and dispel some of the fear of the unknown about the testing situation without compromising any "real tests." In response to this need, The Psychological Corporation has been developing a presentation on tape, accompanied by a practice booklet,[2] to acquaint students and others who might soon be applying for jobs with some of the common types of employment tests. If this approach proves useful, schools and employment offices may find it an economical way to eliminate one source of concern.

As a means of reducing the role of the live examiner in the testing situation, the use of recorded directions to administer actual tests, with both the reading of directions and the timing included on the tape or record, has been suggested. The advantages of such uniform administration are obvious. Moreover, the use of tapes may be more acceptable to those who tend to feel uncomfortable when the examiner is a member of a different group.

Diminishing Cultural Bias

Either because of their content or style, there are some types of test questions which provoke charges of cultural bias from well-meaning persons even when such

[1] New York: Harper & Bros., 1961. Pp. 48-49.

[2] *Test Orientation Procedure (TOP).* See page 8.

charges may have little basis in fact. If it is possible to achieve the purpose of the testing by using measures which do not contain such provocative items, it is manifestly desirable to do so, regardless of the merits of the charges. Many of those who feel that existing tests are too loaded with "middle-class" items to be fair to disadvantaged groups propose "culture-free" or "culture-fair" tests as substitutes. The term "culture-free" is misleading; no instrument which measures behavior can be free of cultural influences. What is sought are "culture-fair" or "cross-cultural" or "culture-common" tests which are measures based on experiences equally familiar or unfamiliar to advantaged and disadvantaged groups. It is often suggested that verbal tests should be replaced by nonverbal measures in order to eliminate cultural bias. However, the preponderant weight of research evidence indicates that nonverbal tests do not measurably benefit disadvantaged groups. In most instances, the disadvantaged score no better on nonverbal or so-called "culture-fair" tests than they do on conventional tests.

Perhaps the answer lies in trying to obtain the "culture-laden" rather than the "culture-free" test. The purpose of employment testing is to select people who ultimately will be successful in one or more jobs. The jobs are inevitably embedded in some cultural matrix, and the criteria of success will undoubtedly be influenced by cultural factors. Thus the abilities to understand oral and perhaps written instructions, to go from one place to another in a reasonable manner, and to cope with simple arithmetic, are activities which are "culture-laden" but which are also likely to be criterion-related. Such behaviors have their parallels in everyday living and can be translated into test questions which are not unfair to disadvantaged applicants.

Prompted by such considerations, a series of tests entitled *Fundamental Achievement Series* is being developed by The Psychological Corporation. The series includes verbal and numerical tests, with many of the items in each test based on "culture-laden" experiences assumed to be quite common in the population.

Investigation of Bonuses and Separate Norms

To compensate the disadvantaged for the effects of deprivation they have experienced, various devices have been proposed to adjust their test scores. One suggestion is that raw score points be given as a bonus, a procedure apparently similar, though not truly comparable, to the awarding of bonus points to veterans taking a civil service examination. Aside from the problems of determining who should be given a bonus and how much bonus, making the award is a discriminatory act which adds nothing to the essential predictive value of the test.

The position that test scores of the disadvantaged should be compared only with the scores of others similarly disadvantaged can become, in some circumstances, a special case of awarding a bonus. Norms based on the disadvantaged provide useful descriptive information and, together with norms for the majority group, the data could be helpful to counselors and psychologists who are advising individuals. But if a separate cutoff score is established for each norms group, in order to yield the same percentage of acceptable cases from the different groups, we have essentially the situation of giving bonus points to the lower-scoring groups.

These approaches to the problem of employing the disadvantaged have little value in selection. Adoption of such procedures would place a burden on the civilian employer which is not undertaken by either the civil service or the military agencies of the United States. It is clearly the obligation of an employer not to discriminate among persons on the grounds of race, religion, or national origin, but it is clearly not the obligation of an employer to hire or to promote the less qualified in an attempt to compensate for some injustice of society in general.

When there is adequate evidence that test scores have different meanings for different groups, an improvement in accuracy of prediction may result from the use of weightings which would be optimum for each group. Experience to date has not revealed frequent occasions for such differential weighting.

Appropriate Use of Test Scores

All of us recognize the importance of the training or apprenticeship period in developing the skills needed for certain kinds of jobs. Some have maintained that during the training period conducted by a company for newly hired people, the disadvantaged can be taught what they need to know, regardless of their scores on the employment tests. Therefore, it is argued, the administration of tests which include questions that have little to do with the job is unjustified. This argument deserves serious examination. Certainly, skills and knowledge can be imparted to motivated people during a training period, and efforts in this direction should be encouraged. Usually such training is more successful for relatively simple jobs which require few skills and operations than for higher level jobs which call for a broader spectrum of previously developed talents. But even in training situations, tests will ordinarily predict the ability to learn

about the job while on the job, or the ability to complete the necessary training. Many companies would like to hire people who can be promoted to better jobs after they have mastered the entry job. Tests which measure general ability—as reflected in verbal and numerical aptitude—are helpful in identifying applicants who would be suitable for promotion in the future. It is important to remember that tests which seem unrelated to immediate jobs might be useful in identifying those who could advance to higher levels.

Tests, like other predictive instruments, must be evaluated in terms of how well predictions made from them conform with reality as represented by measures of success on the job (criteria). Without an adequate criterion, meaningful and free of bias, it will be difficult for anyone to know what predictors are effective. It is obvious to professional users of tests that criteria must be studied and refined; others may occasionally need a reminder not to forget the criteria while they are busy attacking the predictors.

Summing Up

Companies have been accused of using tests as a means of maintaining unfair discrimination against groups which have already suffered from many forms of discrimination. Such practices are a distortion of the proper function of tests and deserve condemnation. It does not follow, however, that tests themselves merit condemnation. Most users in industry expect the tests to help them identify the people best suited for the jobs to be filled.

From the standpoint of corporate management, the employment function can legitimately be viewed as a type of purchasing operation. It then is the duty of the employment manager to hire those candidates who offer the best promise of contributing to the success of the enterprise. The reason that tests have been used for many years by many employers is that, in management's opinion, the information furnished by tests is valuable in making hiring decisions. If the employer sets his minimum scores too high he does not fill the available jobs. If he sets his minimum scores too low he hires persons difficult to train, low in productivity, and high in liability to error. Sophisticated personnel officers realize that predictors such as interviews, reference investigations, and tests are part of a total evaluation problem, that of obtaining better workers or of matching workers and jobs more precisely.

Whenever the number of applicants exceeds the number of job openings, some applicants will be rejected. This is one of the hard facts of life. It is not surprising that those who are rejected sometimes attack the selection procedures on the grounds that these are invalid or unfair. Although testing is not free from defect or beyond criticism, appropriately chosen and properly administered ability tests are superior to most available alternatives.

Some of the problems which stem from the testing of disadvantaged groups are of a technical nature and are not related to the issue of discrimination. Reduced reliability is often due to the fact that score distributions obtained from disadvantaged groups are compressed. The same compression (and reduced reliability) is also found when only highly capable individuals are being tested. These problems require technical study and, in some instances, may call for the use of different measures or the development of new and more appropriate tests. The basic issue is not necessarily one of discrimination against a particular subgroup.

Many of the issues in the testing of disadvantaged groups have both psychometric and social aspects. Some of the current testing procedures should be changed to reduce the fear and hostility that may be engendered by materials felt to be biased or unfair. More attention must be devoted to research with actual score and performance data in order to improve the predictive efficiency of tests. Such changes in employment and research procedures will increase the likelihood of employment tests fulfilling their primary mission of helping the employer select the people best able to do the jobs. These approaches are basically psychometric steps.

Discrimination against disadvantaged groups, which is at the root of the concern of many who attack tests, will not be resolved by improved psychometrics alone. Discrimination in the world of work is a social ailment. Although poor showings on tests may be a symptom of the ailment, the use of tests in employee selection is inherently a friendly rather than a hostile act to those who come to the job market from backgrounds of limited opportunity. Society may well have the responsibility of providing effective remedial instruction for those who have been culturally deprived. The rejection of measuring instruments which register the consequence of such deprivation is merely a modern version of killing the messenger who brings bad news.